System Identification

Arun K. Tangirala

Department of Chemical Engineering
IIT Madras

July 27, 2013

Module 3

Lecture 7

ARMA representations

The MA and AR representations can both be combined under one umbrella

The white-noise $e[k]$ is both the unpredictable component of $v[k]$ and a fictitious input that generates $v[k]$. No stationary process can be expected to be purely either AR or MA type.

A more general representation is one containing mixed effects, which is the auto-regressive moving-average (ARMA(P, M)) representation

$$\boxed{H(q^{-1}) = \frac{C(q^{-1})}{D(q^{-1})} = \frac{1 + \sum_{i=1}^{M} c_i q^{-i}}{1 + \sum_{j=1}^{P} d_j q^{-j}}}$$

On ARMA representations

ARMA models that are both invertible and stationary are allowable representations for causal, stationary processes

- ARMA processes simplify to AR or MA models depending on whether $C(q^{-1}) = 1$ and $D(q^{-1}) = 1$ respectively
- An ARMA process can always be given an equivalent AR or MA representation. In either case, the resulting models are of infinite-order
- **Neither the ACF nor PACF can provide a practically useful distinct signature to determine the order of the numerator and denominator polynomials**
- ARMA model orders have to be determined by trial and error

ACF of an ARMA process

There are primarily two different ways in which one can calculate the theoretical ACF of a ARMA(P, M) process:

1. Convert the given ARMA process to a MA process and use existing results
   - Write the difference equation in MA series (will be of infinite order)
2. Write the difference equations for the ACF
   - Due to the presence of AR and MA characteristics, one obtains two sets of equations
   - The first K set of equations for lags $l = 0, \cdots, K-1$ where $K = \max(P, M+1)$ are linear in the ACVFs. They have to be solved simultaneously
   - The second set of equations are the recursive relations like those that arise in AR process. These are solved recursively for lags $l = K, K+1, \cdots$

Choice of representation

A commonly encountered question is: which form to choose? - MA or AR or ARMA?

- There is no definitive answer since all forms are inter-convertible. A few guidelines, however, exist primarily from an estimation point of view

- No model is a true description of the process!
- A model only allows a convenient way of describing the evolution of a process with time (or space, frequency, etc.)
- Keep the model as simple as and whenever possible (principle of parsimony)

  > Choose that model description which contains few parameters while still maintaining the quality and predictive ability of the model

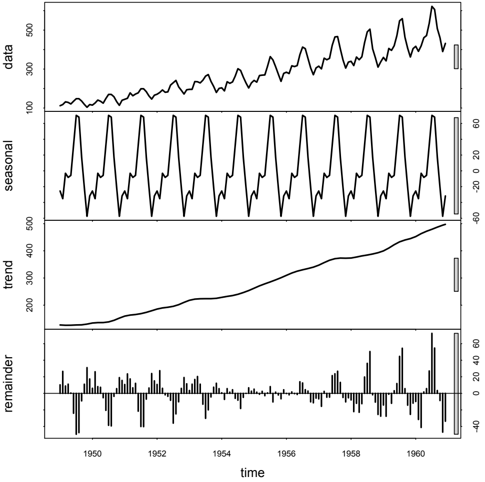
- Over-parametrized model implies more parameters to estimate and hence reliability of the estimates goes down (we give a theoretical proof later)

Non-stationarities

ARMA models are capable of representing almost all classes of (linear) stationary processes. On the other hand, there are several processes that exhibit non-stationarities primarily due to two factors:

1. Trends (polynomial trends)
2. Integrating effects (random walk behaviour)

In addition, the series may contain periodicities and seasonalities. In the discussions to follow we restrict our discussions to non-stationary series only.



Handling non-stationarities

There are primarily two ways of handling non-stationarities:

1. Eliminating trends by first fitting polynomial models to the series and working with the residuals
2. Including integrating effects in the model by suitable differencing of the data

It turns out the second class of methods encompasses the first method whenever

---

If the non-stationarities are due to external signals (exogenous effects), then those effects have to be removed first by fitting a deterministic model between the exogenous signals and then fitting a time-series model to the residuals. Else the deterministic and time-series models can be jointly estimated as well. **This is the crux of system identification**.

Fitting trends

Trends in series can be eliminated by two different means

1. By fitting a trend using polynomial fits
2. Estimating the trend by the use of suitable filters

Consider developing a model for the following series

$$w[k] = \underbrace{\alpha_0 + \alpha_1 k}_{m[k]} + v[k] \quad \alpha_0, \alpha_1 \in \mathcal{R}$$

where $v[k]$ is a zero-mean stationary process

Then, the linear trend (in time) can be estimated by fitting a straight line using a least squares method.

Alternatively,

$$E(w[k]) = E(\beta_0 + \beta_1 k + v[k]) + E(v[k]) = \alpha_0 + \alpha_1 k$$

Thus, the deterministic component (trend) is the average of the series $w[k]$. To put this observation into practice, we replace the theoretical average with an estimate.

Estimating trends using filters

1. Smoothing with a **two-sided finite moving average filter**

$$\hat{m}[k] = \frac{1}{2M+1}\sum_{j=-M}^{M} v[k-j] \tag{63}$$

- This moving average filter assumes linear trend over the interval $[k-M, k+M]$ and that the average of the remaining terms is close to zero
- The filter provides us, therefore, with an estimate of the linear trend

2. A "clever" choice of filter can be used to eliminate polynomial trends. The Spencer 15-point MA filter can be used to estimate polynomial trends of degree 3. See Brockwell [2002] for details.

Estimating the trend by filtering

1. **Exponential smoothing of data to estimate the trend $m[k]$**

$$\begin{aligned} \hat{m}[k] &= \alpha w[k] + (1-\alpha)\hat{m}[k-1], \quad k = 2, \cdots, n \\ \hat{m}[1] &= w[1] \end{aligned}$$

   The choice of α has to be fine tuned on the type of trend that needs to be estimated

2. **Smoothing by elimination of high-frequency components of the series**
   - A spectral analysis of the data can be carried out to determine the cut-off frequency

Note: When seasonality/periodicity is present, an additional procedure has to be followed that eliminates seasonality prior to trend estimation.

Method of differencing

Consider the same series as earlier

$$w[k] = \underbrace{\alpha_0 + \alpha_1 k}_{m[k]} + v[k] \quad \alpha_0, \alpha_1 \in \mathcal{R}$$

Construct the differenced series

$$w[k] - w[k-1] = (1 - q^{-1})w[k] = \beta_1 + v[k]$$

Introducing $\nabla = 1 - q^{-1}$, we can thus observe that the differenced series $\nabla w[k]$ is a non-zero mean stationary process

It is easy to observe that polynomials of d^{th} degree can be eliminated by differencing d times, i.e.,

$$\boxed{\text{Differenced series } \nabla^d w[k] \text{ is free of all polynomial trends upto degree } d}$$

Note: The operator ∇^d should not be confused with the operator $\nabla_d = 1 - q^{-d}$ that is often used to eliminate seasonal effects

Random walk processes

The method of differencing is also capable of handling a wide-class of non-stationarities that are characterized by **random walk** behaviour

The simplest random walk process is

$$v[k] = v[k-1] + e[k] \tag{64}$$

where $e[k]$ is the usual GWN. The adjacent figure shows a sample series from a random walk process.

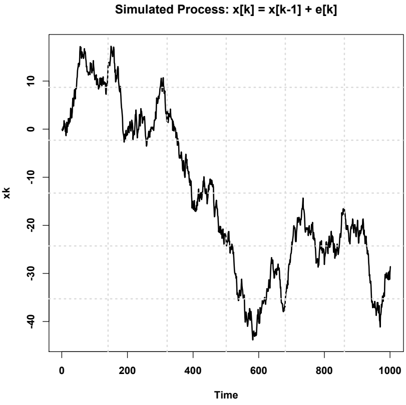


Figure : Sample series from a random walk process

Random walk process ...contd.

The random walk process in (64) which is also an AR(1) process with the pole on the unit circle. It is clearly non-stationary.

An alternative representation of (64)

$$v[k] = \sum_{j=0}^{\infty} e[k-j]$$

Thus, this is an MA process of infinite-order. The coefficients of this model, expectedly, do not satisfy the condition (44)

Note that the differenced series $\nabla v[k]$ is stationary.

ARIMA models

The method of differencing followed by an ARMA representation of the differenced series gives rise to the notion of an ARIMA(P, d, M) model:

$$\boxed{H(q^{-1}) = \frac{C(q^{-1})}{(1-q^{-1})^d D(q^{-1})} = \frac{1 + \sum_{i=1}^{M} c_i q^{-i}}{(1-q^{-1})^d (1 + \sum_{j=1}^{P} d_j q^{-j})}}$$

- The quantities P and M have their usual meaning as in ARMA models
- The parameter d refers to the order of integrating effect
- An ARIMA(P, d, M) model is best remembered as an ARMA model on a series that is differenced d times
- ARIMA models necessarily have d poles on the unit circle

Method of differencing vs. polynomial fits

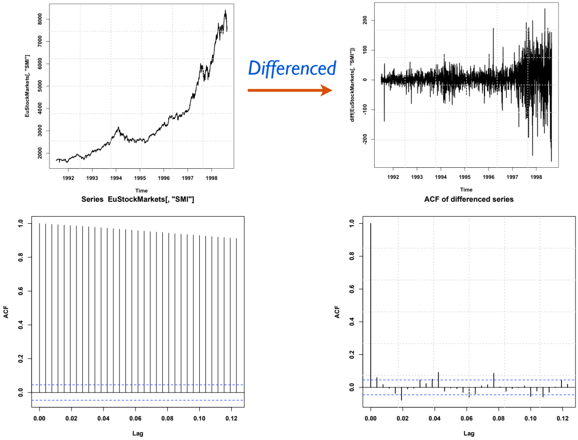
It is useful to observe certain marked differences between the filtering approach and the differencing approach of handling trend-type non-stationarities

- The filtering approach relies on estimation of trend followed by the development of an ARMA model for the stationary residuals. The differencing approach, in contrast, eliminates the trend implicitly and fits an ARMA model.
- Another difference is that an ARIMA model fixes one of the poles of the model to the unit circle irrespective of whether the actual process has a pole on the unit circle or not. This has its own merits and demerits:
  - The merit is that the estimation of "stationary" poles, but close to unit circle, can result in models with confidence regions containing non-stationary models. By forcing the pole to the unit circle a priori, this situation is avoided.
  - The demerits are however that a stationary process with slowly decaying process acquires a non-stationary representation.
  - In system identification, differencing has to be carried out with caution since it amplifies the noise levels in the differenced data thereby decreasing the SNR.

Detecting integrating effects

How does one detect the presence of integrating effects in a time-series?

If the process is purely an integrator or very slowly decaying regressive behaviour, then its ACF estimate never dies out (or rather dies out at very large lags)



- The European Stock Market Index 1992-1998 is shown
- The series is non-stationary
- The ACF shows integrating effects
- The differenced series is more or less stationary, but with a portion of it possessing significantly different variance from the other.

Summary

The contents of this lecture can be summarized as:

- ARMA models are suitable representations for a wide range of stationary processes
- ACF and PACF of ARMA models do not exhibit any distinct signature that can determine the order of an ARMA model
- In choosing a model for a time-series, the principle of parsimony provides a suitable guideline for selecting the appropriate model structure.
- Non-stationary processes can be suitably handled by
    1. Explicitly fitting a deterministic component (either through polynomial fit or by applying a suitable filter) followed by an ARMA modelling of the residual
    2. Fitting ARIMA models, which is equivalent to fitting an ARMA model to the differenced series
- Integrating effects (or slowly decaying nature of stationary processes) are detected by slowly decaying ACFs